

## Cloning and Functional Analysis of the Ubiquitin-specific Protease Gene *UBP1* of *Saccharomyces cerevisiae*\*

(Received for publication, February 5, 1991)

John W. Tobias and Alexander Varshavsky

From the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

In eukaryotes, both natural and engineered fusions of ubiquitin to itself or other proteins are cleaved by processing proteases after the last (Gly<sup>76</sup>) residue of ubiquitin. Using the method of sib selection, and taking advantage of the fact that bacteria such as *Escherichia coli* lack ubiquitin-specific enzymes, we have cloned a gene, named *UBP1*, of the yeast *Saccharomyces cerevisiae* that encodes a ubiquitin-specific processing protease. With the exception of polyubiquitin, the *UBP1* protease cleaves at the carboxyl terminus of the ubiquitin moiety in natural and engineered fusions irrespective of their size or the presence of an amino-terminal ubiquitin extension. These properties of *UBP1* distinguish it from the previously cloned yeast protease *YUH1*, which deubiquitinates relatively short ubiquitin fusions but is virtually inactive with longer fusions such as ubiquitin- $\beta$ -galactosidase. The amino acid sequence of the 809-residue *UBP1* lacks significant similarities to other known proteins, including the 236-residue *YUH1* protease. Null *ubp1* mutants are viable, and retain the ability to deubiquitinate ubiquitin- $\beta$ -galactosidase, indicating that the family of ubiquitin-specific proteases in yeast is not limited to *UBP1* and *YUH1*.

Ubiquitin (Ub),<sup>1</sup> a highly conserved 76-residue protein, is present in eukaryotic cells either free or covalently joined to a great variety of proteins. The posttranslational coupling of ubiquitin to other proteins is catalyzed by a family of Ub-conjugating enzymes (also called E2 enzymes (Pickart and Rose, 1985)) and involves formation of an isopeptide bond between the carboxyl-terminal glycine residue of ubiquitin and the  $\epsilon$ -amino group of a lysine residue in an acceptor protein (Pickart, 1988; Jentsch *et al.*, 1990). A major function of ubiquitin is to mark proteins destined for selective degradation (reviewed by Hershko, 1988; Finley *et al.*, 1988; Ciechanover and Schwartz, 1989; Olson and Dice, 1989; Monia *et al.*, 1990). Ubiquitin has also been shown to have a chap-

erone function, in that its transient (cotranslational) covalent association with specific ribosomal proteins promotes the assembly of ribosomal subunits (Finley *et al.*, 1989). *In vivo* acceptors of ubiquitin include chromosomal histones (Bonner *et al.*, 1988), integral membrane proteins such as the lymphocyte homing receptor and the growth hormone receptor (Siegelman *et al.*, 1986; Leung *et al.*, 1987), actin (Ball *et al.*, 1987), the intracellular neurofibrillary tangles in a variety of neurodegenerative diseases (Perry *et al.*, 1987; Lowe *et al.*, 1988), the plant regulatory protein phytochrome (Jabben *et al.*, 1989), the short-lived MAT $\alpha$ 2 repressor of the yeast *Saccharomyces cerevisiae* (Hochstrasser *et al.*, 1991), and cyclins, a family of short-lived cell cycle regulators (Glötzer *et al.*, 1991).

Unlike branched Ub-protein conjugates, which are formed posttranslationally, linear Ub-protein adducts are formed as the translational products of either natural or engineered ubiquitin gene fusions. Thus, in *S. cerevisiae* for example, ubiquitin is generated exclusively by proteolytic processing of precursors in which ubiquitin is joined either to itself, as in the linear polyubiquitin protein UBI4, or to unrelated amino acid sequences, as in the hybrid proteins UBI1-UBI3 (Özkaynak *et al.*, 1987). In growing yeast cells, ubiquitin is generated from the UBI1-UBI3 precursors whose "tails" are specific ribosomal proteins (Finley *et al.*, 1989). The polyubiquitin (*UBI4*) gene is dispensable in growing cells but becomes essential (as the main supplier of ubiquitin) during stress (Finley *et al.*, 1987, 1988). The lack of genes encoding mature ubiquitin and the fusion structure of ubiquitin precursors in yeast are characteristic of other eukaryotes as well (Schlesinger and Bond, 1987). Thus, there must exist Ub-specific processing proteases that produce ubiquitin, a protein essential for cell viability (Finley *et al.*, 1984, 1987, 1988), by cleaving Gly-X peptide bonds at the junctions between ubiquitin and its carboxyl-terminal extensions. *A priori*, these proteases may also have an isopeptidase activity, *i.e.* the ability to deubiquitinate posttranslationally formed, branched Ub-protein conjugates. A multiubiquitin chain is one such conjugate, in which ubiquitin itself serves as an acceptor, with several ubiquitin moieties attached sequentially to the initial acceptor protein to form a chain of branched Ub-Ub adducts (Chau *et al.*, 1989). Formation of a multiubiquitin chain on a targeted protein is required for the protein's subsequent degradation (Chau *et al.*, 1989; Gregori *et al.*, 1990).

Ub-specific, ATP-independent proteases capable of cleaving ubiquitin from its linear or branched conjugates have been detected in all eukaryotes examined (Matsui *et al.*, 1982; Rose, 1988; Mayer and Wilkinson, 1989; Jonnalagadda *et al.*, 1989) but not in bacteria such as *Escherichia coli*, which lack eukaryotic ubiquitin and Ub-specific enzymes (Finley *et al.*, 1988). Recently, Miller *et al.* (1989) have cloned a *S. cerevisiae* gene, named *YUH1*, for a Ub-specific protease that cleaves ubiquitin from its carboxyl-terminal fusions to relatively short

\* This work was supported by Grants AG08991 and DK39520 (to A. V.) from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M63484.

<sup>1</sup> The abbreviations used are: Ub, ubiquitin; DHFR, dihydrofolate reductase; UB-X-DHFR, ubiquitin-X-mouse dihydrofolate reductase, where X is an amino acid residue at the Ub-DHFR junction; Ub-X- $\beta$ gal, ubiquitin-X-*E. coli*  $\beta$ -galactosidase; MTX, methotrexate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; DTT, dithiothreitol; LB, Luria broth; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; kb, kilobase(s); bp, base pair(s); ORF, open reading frame.

amino acid sequences but is virtually inactive with larger fusions such as Ub- $\beta$ -galactosidase (Ub- $\beta$ gal). Wilkinson *et al.* (1989) have also cloned a cDNA for a mammalian homolog of the yeast YUH1 protease.

We now report the cloning and functional analysis of an *S. cerevisiae* gene, named *UBP1*, which encodes a Ub-specific processing protease whose amino acid sequence is dissimilar to those of the YUH1 protease and other known proteins. The *UBP1* protease cleaves at the carboxyl terminus of the ubiquitin moiety in linear fusions irrespective of their size or the presence of an amino-terminal ubiquitin extension. We also show that null *ubp1* mutants are viable and have properties indicating that the family of Ub-specific proteases in yeast is not limited to *UBP1* and *YUH1*.

Since Ub-specific proteases cleave ubiquitin from its carboxyl-terminal extensions irrespective of the identity of the extension's residue abutting the cleavage site (Bachmair *et al.*, 1986; Gonda *et al.*, 1989), *UBP1* and its functional counterparts make possible the *in vivo* or *in vitro* generation of proteins (including short polypeptides) bearing predetermined residues at their amino termini, a method with applications in both basic research (Baker and Varshavsky, 1991; Bartel *et al.*, 1990; Johnson *et al.*, 1990; Townsend *et al.*, 1988) and biotechnology (Lu *et al.*, 1990; Butt *et al.*, 1989; Sabin *et al.*, 1989; Ecker *et al.*, 1989; Yoo *et al.*, 1989; Mak *et al.*, 1989).

#### EXPERIMENTAL PROCEDURES

**Strains, Media, and Genetic Techniques**—*E. coli* strains HB101 (*pro leu thi lacY hsdS<sub>B</sub>20 endA secA rpsL20 ara-14 galK2 xyl-5 mtl-1 supE44 merB<sub>B</sub>*; Boyer and Roulland-Dussoix, 1969) and JM101 (*supE thi  $\Delta$ (lac-proAB)/F' tra $\Delta$ 36 proA<sup>+</sup> proB<sup>+</sup> lacI<sup>q</sup> lacZ $\Delta$ M15*; Yanisch-Perron *et al.*, 1985) were propagated in Luria broth (LB) and used as hosts for plasmids and phage M13 derivatives. The *S. cerevisiae* strains used were DF5 (*MATa/MATa trp1-1/trp1-1 ura3-51/ura3-51 his3- $\Delta$ 2000/his3- $\Delta$ 2000 leu2-3,112/leu2-3,112 lys2-801/lys2-801 gal/gal*; Finley *et al.*, 1987), BBY45 (*MATa trp1-1 ura3-52 his3- $\Delta$ 2000 leu2-3,112 lys2-801 gal*; Bartel *et al.*, 1990), and BWG1-7a (*MATa his4 ura3 ade1 leu2*; Guarente *et al.*, 1982). Rich (YPD) and synthetic yeast media were prepared according to Sherman *et al.* (1986). Transformations of *S. cerevisiae* and *E. coli* were carried out, respectively, by the lithium acetate method (Ito *et al.*, 1983) and by the calcium chloride method (Ausubel *et al.*, 1989). The DNA library RB237 used for the isolation of *UBP1* was obtained by cloning a partial *Sau3AI* digest of *S. cerevisiae* genomic DNA into the *Bam*HI site of the *URA3-CEN4*-based vector *YCp50* (Rose *et al.*, 1987).

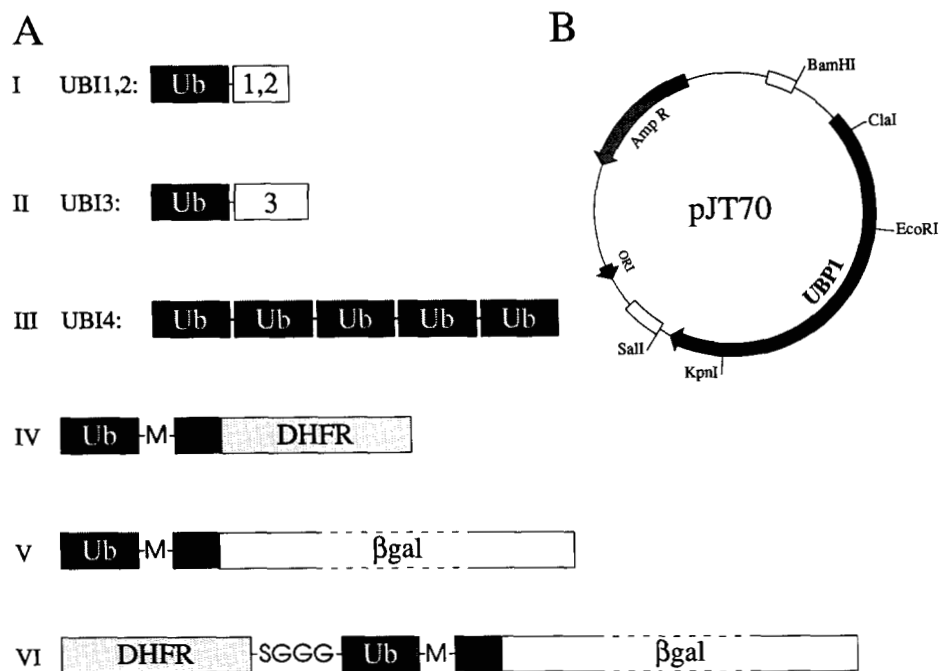
**Ubiquitin-containing Test Proteins**—Ub-M- $\beta$ gal (ubiquitin-Met- $\beta$ -galactosidase) (see Fig. 1A, V) and the otherwise identical Ub-P- $\beta$ gal containing proline (instead of methionine) at the Ub- $\beta$ gal junction were purified by affinity chromatography of extracts from [<sup>35</sup>S]methionine-labeled *E. coli* (carrying pKKUb-M- $\beta$ gal or pKKUb-P- $\beta$ gal) on aminophenylthiopyranogalactoside-Sepharose as described by Gonda *et al.* (1989). Ub-M-DHFR (ubiquitin-Met-mouse dihydrofolate reductase) (Fig. 1A, IV) was prepared as follows: an overnight culture (1 ml) of *E. coli* JM101 carrying the plasmid pUb-M-DHFR (Bachmair and Varshavsky, 1989) was diluted into 50 ml of LB supplemented with ampicillin at 40  $\mu$ g/ml, and the cells were grown with shaking at 37 °C to an A<sub>600</sub> of ~0.9. The culture was chilled on ice for 15 min, centrifuged at 3000  $\times$  g for 5 min, and washed twice with M9 buffer (Ausubel *et al.*, 1989) at 4 °C. The cells were resuspended in M9 buffer supplemented with glucose (0.2%), thiamine (2  $\mu$ g/ml), ampicillin (40  $\mu$ g/ml), 1 mM isopropylthiogalactoside, and methionine assay medium (0.0625%) (Difco). After incubation with shaking for 1 h at 37 °C, 1 mCi of Trans<sup>35</sup>S-label (ICN; at least 70% of <sup>35</sup>S is in L-methionine, with the rest in L-cysteine and related compounds) was added, and shaking was continued for 5 min. Unlabeled L-methionine was then added to 1 mM, and shaking was continued for another 10 min. Cells were harvested, washed twice with M9 buffer, and resuspended in 0.5 ml of 25% (w/v) sucrose, 50 mM Tris-HCl, pH 8.0. 0.1 ml of freshly prepared egg-white lysozyme solution (10 mg/ml; Sigma) in 0.25 M Tris-HCl, pH 8.0, was then added, and the mixture was incubated for 5 min at 0 °C, followed by the addition of 0.1 ml of 0.5 M Na-EDTA, pH 8.0, and further incubation at 0 °C for 5 min. The suspension was transferred into a

centrifuge tube containing 0.975 ml of 65 mM Na-EDTA, 50 mM Tris-HCl (pH 8.0), and protease inhibitors antipain, chymostatin, leupeptin, pepstatin, and aprotinin, each at 25  $\mu$ g/ml. 10  $\mu$ l of 10% (v/v) Triton X-100 was then added and dispersed by pipetting. The lysate was centrifuged at 40,000  $\times$  g for 30 min. The supernatant, containing the <sup>35</sup>S-labeled Ub-M-DHFR, was frozen in liquid nitrogen and stored at -85 °C. To purify Ub-M-DHFR, a methothrexate (MTX) affinity matrix was prepared as described by Kaufman (1974), using amino-hexyl-Sepharose (Sigma). The MTX-Sepharose column (0.5-ml bed volume) was washed with 10 ml of MTX buffer (0.2 M NaCl, 1 mM Na-EDTA, 0.2 mM dithiothreitol (DTT), 20 mM Na-HEPES, pH 7.5). The Ub-M-DHFR-containing <sup>35</sup>S supernatant was thawed, clarified by centrifugation at 12,000  $\times$  g for 1 min, and applied to the column, which was then washed with 50 ml of MTX buffer, 50 ml of MTX buffer containing 2 M urea, and then again with 50 ml of MTX buffer. Ub-M-DHFR was then eluted from the column with folic acid buffer (10 mM folic acid, 1 M KCl, 1 mM DTT, 1 mM Na-EDTA, 0.2 M potassium borate, pH 9.0). The folic acid buffer was applied in 1-ml aliquots, and 1-ml fractions were collected. The fractions were assayed for <sup>35</sup>S; pooled fractions containing the peak of <sup>35</sup>S were dialyzed for 20 h at 4 °C against two changes of storage buffer (50% glycerol, 1 mM MgCl<sub>2</sub>, 0.1 mM Na-EDTA, 40 mM Na-HEPES, pH 7.5), and stored at -20 °C. The purified [<sup>35</sup>S]Ub-M-DHFR was assayed by SDS-PAGE and fluorography and found to be greater than 95% pure. Other test proteins such as DHFR-Ub-M- $\beta$ gal (Fig. 1A, VI), and the natural yeast ubiquitin fusions UB12 (Fig. 1A, I), UB13 (Fig. 1A, II), and UB14 (polyubiquitin; Fig. 1A, III), were expressed in *E. coli* and assayed for deubiquitination with *UBP1* in mixed *E. coli* extracts as described below.

**Sib Selection Cloning of the *UBP1* Gene**—*E. coli* HB101 cells carrying the *S. cerevisiae* genomic DNA library RB237 (see above) were spread on LB plates containing 100  $\mu$ g/ml ampicillin (LB/amp) to yield ~40 colonies per plate. After ~16 h at 36 °C, the colonies were replica-plated onto LB/amp plates; the original plates were stored at 4 °C, while their replicas were incubated at 36 °C for 24 h. Cells on each of the replica plates were then eluted with 1 ml of LB/amp by repeated washing over the agar surface until all of the colonies were dispersed into the liquid. The eluates were each diluted with 4 ml of LB/amp and incubated overnight with shaking at 36 °C. 1.7 ml of each overnight culture was chilled on ice and centrifuged at 12,000  $\times$  g for 1 min. The pellet was resuspended by vortexing in 50  $\mu$ l of 25% (w/v) sucrose, 0.25 M Tris-HCl (pH 8.0). 10  $\mu$ l of freshly prepared egg-white lysozyme solution (10 mg/ml; Sigma) in 0.25 M Tris-HCl (pH 8.0) was added, mixed by light vortexing, and incubated at 0 °C for 5 min. 0.15 ml of 75 mM Na-EDTA, 0.33 M Tris-HCl (pH 8.0) was then added, followed by incubation at 0 °C for 5 min, with occasional stirring. 1  $\mu$ l of 10% (v/v) Triton X-100 was then added and mixed by pipetting. Each of the resulting lysates was centrifuged at 12,000  $\times$  g for 15 min at 4 °C, and the supernatants were tested for Ub-specific protease activity as described below, using Ub-M-DHFR as the test substrate. Upon finding a positive extract, overnight cultures were made as described above from each of the ~40 individual colonies on the plate that yielded the positive extract. A *YCp50*-based plasmid, named pJT55, was then isolated from a transformant that yielded a positive extract upon the second round of screening. The yeast DNA insert of pJT55 was analyzed as described below.

**In Vitro Assays for Ub-specific Protease Activity**—An *E. coli* extract (9  $\mu$ l) prepared as described above was added to 1  $\mu$ l of an <sup>35</sup>S-labeled test protein (Ub-M-DHFR, Ub-M- $\beta$ gal, or Ub-P- $\beta$ gal) in storage buffer (see above), and incubated at 36 °C for 3 h. 5  $\mu$ l of 3-fold concentrated SDS-PAGE sample buffer (6% SDS, 33% glycerol, 30 mM Na-EDTA, 0.3 M DTT, 0.6% bromophenol blue, 0.3 M Tris-HCl, pH 6.8) was then added, followed by heating of the sample at 100 °C for 3 min, electrophoresis in a 6% (for  $\beta$ -galactosidase) or a 12% (for DHFR) polyacrylamide-SDS gel, and fluorography.

Similar tests using the unlabeled natural ubiquitin precursors UB12, UB13, and UB14 (see introduction) were carried out as follows. 5  $\mu$ l of an *E. coli* extract to be tested for Ub-specific protease activity was mixed with 5  $\mu$ l of an *E. coli* extract prepared as described above from *E. coli* carrying a plasmid that expressed *UB12*, *UB13*, or *UB14* (Özkaynak *et al.*, 1987; Finley *et al.*, 1987). The mixture was incubated for 1 h at 36 °C, followed by SDS-PAGE as described above, and immunoblot analysis using a rabbit polyclonal anti-Ub antibody (a gift from Dr. Vincent Chau, Wayne State University, Detroit MI). Specifically, after SDS-PAGE, the gel and PVDF membrane (Millipore) were equilibrated with transfer buffer (20% methanol, 0.2 M glycine, 25 mM Tris), followed by electroblotting of proteins from the gel using the Genie apparatus (Ideia Scientific) at 20 V for 25 min.



**FIG. 1. Test ubiquitin fusions and *UBP1*-containing plasmid pJT70.** In *A*: *I*, the yeast *UBI1* and *UBI2* gene products (their amino acid sequences are identical) are the natural fusion of ubiquitin to a specific ribosomal protein (Finley *et al.*, 1989); *II*, the yeast *UBI3* protein is the natural ubiquitin fusion to another ribosomal protein; *III*, the yeast *UBI4* protein is the natural head-to-tail, spacerless fusion of five ubiquitin moieties (Özkaynak *et al.*, 1987); *IV*, Ub-M-DHFR is an engineered fusion of yeast ubiquitin to mouse dihydrofolate reductase (*DHFR*), the junctional residue being methionine (*M*); *V*, same as *IV* but a fusion to *E. coli*  $\beta$ -galactosidase; *VI*, DHFR-Ub-M- $\beta$ gal, same as *V* but with the ubiquitin moiety bearing at its amino terminus the mouse *DHFR* moiety (connected to ubiquitin via a Ser-Gly-Gly-Gly linker). Black rectangles indicate the 45-residue sequence from an internal region of *E. coli lac* repressor that is present at the amino termini of *DHFR* and  $\beta$ gal in these engineered test proteins (Bachmair and Varshavsky, 1989). With the exception of  $\beta$ gal, the lengths of rectangles are proportional to the sizes of protein moieties. (See "Experimental Procedures" for details.) In *B*: pJT70, the plasmid that carries the ~2.8-kb, *UBP1*-containing yeast genomic DNA insert. (See "Experimental Procedures" and text for details.)

The membrane was washed in ANT buffer (0.15 M NaCl, 0.02% NaN<sub>3</sub>, 50 mM Tris-HCl, pH 8.0), then incubated (at room temperature) with rocking in ANT buffer plus 20% fetal calf serum, followed by a 2-h incubation at room temperature with the anti-Ub antibody in ANT/fetal calf serum buffer. The membrane was then washed several times in ANT buffer plus 0.05% Tween 20. An alkaline phosphatase-linked anti-rabbit second antibody (Promega Biotech) in the ANT/Tween buffer was added, followed by a 30-min incubation at room temperature, washing of the membrane in ANT/Tween buffer, and detection of the bound second antibody using a chromogenic phosphatase substrate (Sigma).

A positive control for a Ub-specific processing protease activity was an extract prepared from *S. cerevisiae* (strain BWG1-7a): cells from 1.5 ml of a stationary culture grown in YPD were pelleted by centrifugation, washed twice with H<sub>2</sub>O, and resuspended in 0.15 ml of 1% (v/v) Triton X-100, 1 mM Na-EDTA, 50 mM Na-HEPES (pH 7.5), containing antipain, chymostatin, leupeptin, pepstatin, and aprotinin, each at 25  $\mu$ g/ml, and (freshly added) 1 mM phenylmethylsulfonyl fluoride. 0.3 g of 0.5-mm glass beads was then added, and the cells were disrupted by vortexing at 4°C for 1 min three times, with 1-min pauses on ice. The suspension was centrifuged at 12,000  $\times$  g for 5 min, and the supernatant was used as a positive control in the *in vitro* deubiquitination assays.

**In Vivo Assay for Ub-specific Protease Activity**—The *UBP1* protease activity with one of the test proteins, DHFR-Ub-M- $\beta$ gal, in which ubiquitin bore an amino-terminal and a carboxyl-terminal extension, was assayed using *E. coli* that expressed both *UBP1* and DHFR-Ub-M- $\beta$ gal. A plasmid (p DHFRUbM $\beta$ gal) that expressed the latter protein in *E. coli* from the yeast *GAL10* promoter was constructed by inserting the ~0.67-kb *Bam*HI/*Hind*III fragment from the plasmid pMet-1, which contained the complete mouse *DHFR* coding sequence (Bachmair and Varshavsky, 1989), into the *Kpn*I site-containing version of the pUB23 plasmid (Gonda *et al.*, 1989). Site-directed mutagenesis using the Muta-Gene kit (Bio-Rad) was employed to position the start codon of the *DHFR* reading frame downstream of the *GAL10* promoter of pUB23 (Bachmair *et al.*, 1986)

and to produce a reading frame that encoded a 4-residue linker sequence (Ser-Gly-Gly-Gly) between the carboxyl terminus of *DHFR* and the amino terminus of ubiquitin, yielding the plasmid pDHFRUbM $\beta$ gal.

*E. coli* JM101 carrying pDHFRUbM $\beta$ gal, alone or together with the compatible plasmid pJT184 (see below) that expressed *UBP1*, was grown in LB (in the presence of ampicillin at 40  $\mu$ g/ml and chloramphenicol at 35  $\mu$ g/ml to maintain the plasmids) to an *A*<sub>600</sub> of ~0.3. 1 ml of the culture was centrifuged at 12,000  $\times$  g for 1 min. The cells were resuspended in 50  $\mu$ l of M9 buffer containing glucose (0.2%) and thiamine (2  $\mu$ g/ml) and incubated at 36°C for 10 min, followed by the addition of 20  $\mu$ Ci of Trans<sup>35</sup>S-label (ICN) and a further 2-min incubation. Unlabeled L-methionine was then added to a final concentration of 1 mM, the cells were incubated for 10 min at 36°C, then centrifuged at 12,000  $\times$  g for 30 s, washed once in cold M9 buffer, and resuspended in 50  $\mu$ l of lysis buffer (4% SDS, 125 mM Tris-HCl, pH 6.8), followed immediately by heating at 100°C for 4 min. The sample was then added to 1 ml of immunoprecipitation buffer (1% Triton X-100, 0.5% Na-deoxycholate, 0.15 M NaCl, 5 mM Na-EDTA, 50 mM Na-HEPES, pH 7.5, plus freshly added 1 mM phenylmethylsulfonyl fluoride), centrifuged at 12,000  $\times$  g for 10 min at 4°C, and the upper 0.9 ml of the supernatant was added to 6  $\mu$ l of a concentrated hybridoma supernatant containing a monoclonal antibody to  $\beta$ gal (Bachmair *et al.*, 1986). The sample was incubated in a microcentrifuge tube for 1 h at 0°C. 10  $\mu$ l of protein A-agarose (Repligen) was then added; the suspension was incubated with rocking for 30 min at 4°C, and centrifuged for 15 s at 12,000  $\times$  g. The pellet was washed three times with immunoprecipitation buffer containing 0.1% SDS, dissolved in the SDS-PAGE, sample buffer, and electrophoresed in a 6% polyacrylamide-SDS gel, followed by fluorography.

**Plasmid Manipulations and DNA Sequencing**—Plasmids were constructed using standard procedures (Ausubel *et al.*, 1989). DNA fragments were isolated from agarose gels using GeneClean (Bio101, La Jolla, CA). The YCp50-based plasmid (named pJT55) that conferred deubiquitinating activity onto *E. coli* (see above and "Results") had a

~15-kb yeast genomic DNA insert (data not shown). *SphI* digestion of pJT55 yielded a ~10-kb fragment which, upon subcloning into the *SphI*-cut vector pUC19 (Yanisch-Perron *et al.*, 1985), conferred the same deubiquitinating activity (plasmid pJT57). The ~10-kb yeast DNA insert of pJT57 was cut with *SphI* and *XhoI*, yielding a ~5.5-kb fragment, which was subcloned into pUC19. The resulting plasmid, pJT60, also conferred deubiquitinating activity onto *E. coli* (see "Results").

Since further subcloning of the pJT60 insert, using its single *KpnI* site, did not yield a deubiquitinating activity-conferring fragment, the two fragments of the *KpnI*-cut, ~5.5-kb insert of pJT60 were subcloned into M13mp18 and M13mp19 (Yanisch-Perron *et al.*, 1985), and sets of nested deletions were generated using the 3' exonuclease activity of T4 DNA polymerase (Dale *et al.*, 1985). Single-stranded M13 DNA was sequenced using the Sequenase kit (U. S. Biochemical Corp.) under conditions recommended by the manufacturer. The internal *KpnI* site in pJT60 was found to reside within a 2427-bp open reading frame (ORF), which was sequenced on both strands.

The sequence information was then used to subclone further the pJT60 insert. Following digestion with *AccI*, the 5' overhangs of the resultant ~2.8-kb fragment were filled in with Klenow PolI, followed by ligation of *SalI* linkers. The fragment was then digested with *SalI* and *BamHI*, purified by agarose gel electrophoresis, and subcloned into *SalI/BamHI*-cut pUC19. The resultant plasmid, pJT70, conferred the deubiquitinating activity onto *E. coli* as efficiently as the initial pJT57 plasmid (data not shown), indicating that the corresponding Ub-specific protease, named UBP1, was encoded by the 2427-bp ORF within the ~2.8-kb insert of pJT70 (Fig. 1B).

For experiments that involved the simultaneous presence of two distinct plasmids in *E. coli*, the UBP1-containing insert of pJT70 was excised with *SalI* and *BamHI*, and subcloned into the *SalI/BamHI*-cut vector pACYC184 (New England Biolabs), yielding pJT184. pACYC184 bears the P15A origin of replication, which makes it compatible with pMB1(ColE1)-based *E. coli* vectors such as pUC19 and pBR322 (Sambrook *et al.*, 1989).

The nucleotide and (predicted) amino acid sequences of the UBP1 gene and protein were compared to sequences in the GenBank (version 66.0), EMBL (version 23.0), and GENPEPT (a translated version of GenBank) databases, using the FastA algorithm (Pearson and Lipman, 1988).

**Construction of a Null *ubp1* Mutant**—A *ubp1* null allele was made *in vitro* by replacing 81% of the UBP1 coding region with a fragment carrying the yeast *URA3* gene. First, the entire ~15-kb yeast DNA insert of pJT57 was subcloned into pUC9 to yield pJT9, which was then digested with *ClaI* and *KpnI*. The resultant ~2-kb fragment, encompassing 81% of the UBP1 ORF, was replaced with the *URA3*-containing ~1.1-kb fragment of pRBU1, yielding pJT9K. The plasmid pRBU1, kindly provided by Dr. R. T. Baker, was constructed by isolating the ~1.1-kb *HindIII/SmaI* fragment from YEp24 (Botstein *et al.*, 1979), filling in the *HindIII* site using Klenow PolI, and ligating the blunt-ended fragment into *SmaI*-cut pUC19.

The ~7.7-kb linear DNA from *SalI*-cut pJT9K that included the deletion/disruption *ubp1-Δ1* allele was used to transform both haploid (BBY45) and diploid (DF5) *ura3* yeast strains. Southern hybridization analysis of haploid and diploid *Ura*<sup>+</sup> transformants (with the latter expected to be heterozygous at the UBP1 locus) confirmed the predicted structure of the *ubp1-Δ* allele by showing the presence of restriction endonuclease sites diagnostic of the transplacement (data not shown) (Rothstein, 1983).

## RESULTS

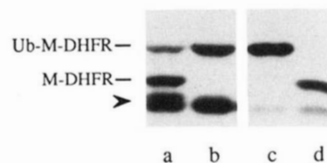
**The UBP1 Gene**—The previously cloned Ub-specific processing protease of *S. cerevisiae*, YUH1, cleaves ubiquitin from relatively short carboxyl-terminal extensions of ubiquitin but is virtually inactive with larger ubiquitin fusions such as Ub-gal (Miller *et al.*, 1989).<sup>2</sup> The efficient deubiquitination of the latter fusion in yeast cells and cell-free extracts (Bachmair *et al.*, 1986; Bachmair and Varshavsky, 1989; see also below), as well as the properties of null *yuh1* mutants (Miller *et al.*, 1989) indicated that *S. cerevisiae* contains a family of Ub-specific processing proteases. To clone the corresponding genes, we took advantage of the fact that bacteria such as *E. coli* lack eukaryotic ubiquitin and Ub-specific enzymes (Finley

*et al.*, 1988). Specifically, the sib selection strategy (McCormick, 1987; Lederberg, 1989) was used to isolate a gene from a *S. cerevisiae* genomic DNA library that could confer Ub-specific protease activity onto *E. coli*. To be clonable by this strategy, a yeast gene must have a promoter that functions in *E. coli* (a minority of yeast promoters are able to do so), must lack introns in coding region (most yeast genes lack introns), and must encode a protein that functions as a monomer or a homooligomer and that is active in *E. coli* extracts.

*E. coli* transformed with a *S. cerevisiae* genomic DNA library (see "Experimental Procedures") were spread on plates to yield ~40 individual colonies per plate. Overnight cultures were then seeded by pooled inoculums from a replica of each plate. Extracts from these cultures were tested for their ability to deubiquitinate Ub-M-DHFR (Fig. 1A IV), and one of the extracts showed the activity (Fig. 2). Overnight cultures were then made as above from each of the ~40 individual colonies on the plate that yielded the positive extract. Extracts from these cultures were tested for the deubiquitinating activity with Ub-Met-DHFR, and a single positive extract was obtained.

A plasmid (named pJT55) was then isolated from the individual *E. coli* transformant that yielded the positive extract. Its ~15-kb yeast DNA insert was subcloned to yield a ~5.5-kb fragment (the plasmid pJT60) that was indistinguishable from the original insert in conferring the deubiquitinating activity onto *E. coli* (see "Experimental Procedures"). Since further subcloning, using the *KpnI* site in the middle of the 5.5-kb insert of pJT60, did not produce activity-conferring fragments, the insert was sequenced in both directions from the *KpnI* site, which was found to reside within an open reading frame (ORF). This sequence information was used to subclone further the insert in pJT60, yielding pJT70 (Fig. 1B), whose ~2.8-kb, activity-conferring insert contained the 2427-bp ORF (Fig. 3) encoding an acidic (calculated pI of 4.63), 809-residue (~93 kDa) protein, which was named UBP1 (for ubiquitin-specific protease). This designation conforms to the existing convention for naming genes encoding components of the ubiquitin system in *S. cerevisiae* (UBI1-UBI4 for the ubiquitin genes (Özkaynak *et al.*, 1984, 1987; Finley *et al.*, 1987, 1988); UBC1-UBC7 for the genes encoding Ub-conjugating enzymes (Jentsch *et al.*, 1987, 1990); UBR1 for the gene encoding the recognition component of the N-end rule pathway, one Ub-dependent proteolytic pathway (Bartel *et al.*, 1990)).

The position of the start (ATG) codon in UBP1 was inferred so as to yield the longest ORF. No ATGs occur in any of the three forward reading frames upstream of this putative



**FIG. 2. Cloning of the UBP1 gene.** A sib selection strategy was used to isolate a gene from a *S. cerevisiae* genomic DNA library that could confer Ub-specific protease activity onto *E. coli* cells that carried such a gene. Lane a, partial *in vitro* deubiquitination of the <sup>35</sup>S-labeled Ub-M-DHFR by an extract from the (pooled) 40 individual *E. coli* transformants produced using the RB297 yeast DNA library (see "Experimental Procedures"). Lane b, same as lane a but an extract from a pool that lacked the deubiquitinating activity. Lane c, purified, <sup>35</sup>S-labeled Ub-M-DHFR (control). Lane d, same as lane c but after incubation with a whole cell yeast extract (see "Experimental Procedures"). Arrowhead denotes an unrelated cleavage product of Ub-M-DHFR that formed in the control (UBP1-lacking) *E. coli* extract.

<sup>2</sup> J. Tobias and A. Varshavsky, unpublished results.



-192	TGTGATCTGGCTCTTTTCTCAGGAAAAAATTTTATAGACATTCAGAAATAGAGCGATTGTCAAATTCGCTCTCTCTCT	
-104	TTTCATTATAAGCTGTGATCATTTTAACTGCTTTCAGTGGCTTCTGTTGCAAACTAGATCTTTTGAACACTCTTCCCTTTTAACT	
-12	ACAAAAATTTGTATGGATTGTTTATTGAAGCAAGATAAACAGTTTATTACAAATTTTATTTGGTTCOCACAGGATTTTGTGAGAAAT	
	M D L F I E S K I N S L L Q F L F G S R Q D F L R N	26
79	TTTAAACTTGGAGTAACAAATAACATCTATCGATTATTTATTAATTTTGGCATAGTAGTATTTTATTAATAAACACAGCCAT	
	F K T W S N N N N N L S I Y L L I F G I V V F F Y K K P D H	56
169	CTAACTACATTTGTAGAGCGTTAGTGAATGACAACAACTTCAGAAATAATAGCGTTAGCGGTGGTTCGCCAGAGTAAGTTT	
	L N Y I V E S V S E M T T N F R N N N S L R N R F S K A G	86
259	ACCCACTTAGAGAGAGATCTTGAAGAGAGTGGTTTCATTTGCTGGTTAGTATTAATGATGTAACACTTGTTTATGAACCTGTGTTT	
	T H L D E E I L K R G G F I A G L V N D G N T C F M N S V L	116
349	CAATCATTTGGCATCATCAGAGATTAATGAGATTCTTGCAATAATGTCAAGACCTATGAGGAGATAGAACAAATGAACAAAT	
	Q S L A S S R E L M E F L D N N V I R T Y E E I E Q N E H N	146
439	GAAGAAGGAACGGGCAAGATTCCTGCTCAAGATGAGCCACTCATAGAAAACACTCGTAAGGCGCAAGAGTTTATGCTGAAGCATAG	
	E E G N G Q E S A Q D E A T H K K N T R K G G K V Y G K H K	176
529	AAGAAATGAATAGGAAGTCAAGTTGGAAG	
	K K L N R K S S S K E D E E K S Q E P D I T F S V A L R D L	206
619	CTTTCTGCTTAAATGCGAAGTATTATCGGGATAACCTTATTCAAAACCAATAGTTTATTTGAAGCAATGTCCAAATCTCCAGAGAAA	
	L S A L N A K Y Y R D K P Y F K T N S L K A M S K S P R K	236
709	AATATTCTTCTGGCTACGACCAAG	
	N I L L G Y D Q E D A Q E F F Q N I L A E L E S N V K S L N	266
799	ACTGAAAACTAGATACACTCCAGTTGCGAATCAGAAATTAACCGATGATGCTTTAGTAGTCAACTTAACTTAACTTAACTTAACTTAA	
	T E K L D T T P V A K S E L P D D A L V G Q L N L G E V G T	296
889	GTTTACATTCAGTGAACAGATGATCTCACTATACATGACAGTCCAAATTCACAACTTCAAACTTAATGACTGCTCT	
	V Y I P T E Q I D P N S I L H D K S I Q N F T P F K L M T P	326
979	TTAGATGATATCAGCGCAGAAAGATTTGTTTACAGTGTGGTGAAGCGGTGGCATATATTCGATTTTCGGGATTAAGCTTA	
	L D G I T A E R I G C L Q C G E N G G I R A Y S V F S G L S	356
1069	AATTTACGAGAGAGATTTGTTTCCACTTTAAATATCTCAGTTATTAAGCGCTGGATGATGATGATGATGATGATGATGATGATGAT	
	N L P N E N I G S T L K L S Q L L S D W S K P E I I E G V E	386
1159	TGTAACCGTTGTGCGCTCACAGCAGCGACTCTCATTTTATTTGGTTCAGTTGAAGAATTTGAAAAAAGCTGAGGGTTCGATCCAGAA	
	C N R C A L T A A H S H L F G Q L K E F E K K P E G S I P E	416
1249	AAGCCAAATTAAGCTGTAAAGATAGAGTCCATCAATTCGAAGAGTCTTGGCAACAGCTTATGACGATGAAGATTATGAAGAGTTG	
	K P I N A V K D R V H Q I E E V L A K P V I D D E D Y K K L	446
1339	CATACAGCAATATGGTACGTAATGCTCAATATCAAGCAGATTTTAATATCAAGACCTCCACATTTATTCATTCATATCAACAGA	
	H T A N M V R K C S K S K Q I L I S R P L I S R P L I S R P L I S R P L I S R P L I S R P L I S R P L I S R P L I S R P	476
1429	TCGATATTGATCCAGAGAGTACATGATTAGAAAAAATACGCAAGATTTGTTTAAAGTCAAGTTGAATCTTGCCCATGATGATGATGAT	
	S V F D P R T Y M I R K N N S K V L F K S R L N L A P W C C	506
1519	GATATTAAATGAATCAATTTGGATGCTGTTTCCCAATGTCAAAAAAGAAAAAGCTGCGCAACAGATTCAAGTGAAGATGAAGACATT	
	D I N E I N L D A R L P M S K K E K A A Q Q D S S E D E N I	536
1609	GGCGGTGATACATATACGAAATTTACATGACGCTTCGAGCAGGAATTTGAAGACAGCAGGAGAAAGATATGATGATGATGATGATGAT	
	G G E Y Y T K L H E R F E Q E F E D S E E E K E Y D D A E G	566
1699	AACATGCGTCTCATTAACATCATACCAAGATATCAGTAATATGATCCCTTAACCGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	
	N Y A S H Y N H T K D I S N Y D P L N G E V D P L N G E V D P L N G E V D P L N G E V D P L N G E V D P L N G E V D P	596
1789	GAT	
	D E Y I E E T D A L G N T I K K R I E I H S D V E N E N V K	626
1879	GATAATGAAGAACTGCAAGAAATCGCAATGTGAGCCTTGACGAACCAAGATCAATGTTGAAGTCAATAGAACATCATCTGATGAG	
	D N E E L Q E I D N V S L D E P K I N V E D Q L E T S S D E	656
1969	GAAGATGTTATACAGCTCCACTCAATGCTAGTTCATTTTCCAGTTCCAGCCACTCCATGACATATTTCATGCGCTCTGCTG	
	E D V I P A P P I N Y A R S F S T V P A T P L T Y S L R S V	686
2059	ATTGTTCACTAGGTAACCAATATGTTGCTCATTCATTCATTTAGAAAAACAGGGGTTGTTGGTGGAGAAATCTGATGAGACTGTG	
	I V H Y G T H N Y G H Y I A F R K Y R G G V T I R S D E T V	716
2149	TACGTTTGGAGAGAGTGAAGTCTTCAACACCGGCTGATTTATGTTATTTTACGAATGATGATTTGATGAAGAACTGGGAGAGATG	
	Y V V D E A E V L S T P G V F M L F Y E Y D F D E E T G K M	746
2239	AAGGATGATTGGAAGCTATTCAGAGTAATTAAGAGATGATGAAAAAGAGCAGGAGCAAAAGAGTCCAGGAGCAAAAGAGAGC	
	K D D L E A I Q S N N E E D D E K E Q E Q K G V Q E P K E S	776
2329	CAAG	
	Q E Q G E G E E Q E E G Q E Q M K F E R T E D H R D I S G K	806
2419	GATGTAAGTAACTGATATTAATACGATATCCGTAATTTGTGTAATAACAATAACTATAATTAATTAATTAATTAATTAATTAATTAAT	
	D V N ter	809
2509	GTTAAATCAATTTTATGCTAGTACGAATGTCTAAAGTTTGTAGGACAAATTCGAAAAACACTTCCATTATTATACAAATCTCTCTA	
2599	AGCTTCATTTTCTTACCATGTACTCTTCAACCTTTTCT	2653

FIG. 3. Nucleotide sequence of the *UBP1* gene and deduced amino acid sequence of the *UBP1* protein.

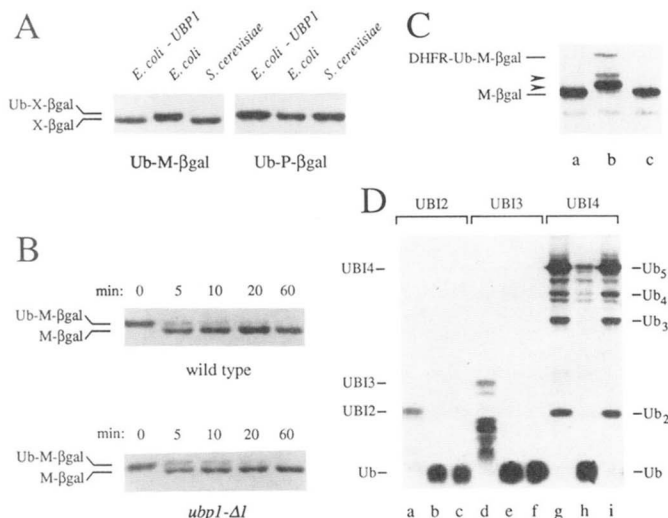
initiator ATG until position -200 (Fig. 3). The codon adaptation index (calculated according to Sharp and Li, 1987) for *UBP1* is 0.19, characteristic of weakly expressed yeast genes. Computer-aided comparisons of the *UBP1* nucleotide sequence and of the predicted amino acid sequence of *UBP1* to sequence databases (see "Experimental Procedures") revealed no significant similarities to sequenced genes or gene products. In particular, *UBP1* lacks sequence similarities to proteins that are either known or expected to interact with ubiquitin, including several yeast Ub-conjugating (UBC) enzymes (Jentsch *et al.*, 1990), UBCs from other organisms (Sullivan and Vierstra, 1989; Reynolds *et al.*, 1990; Schneider *et al.*, 1990), and, in addition, more surprisingly, the YUH1 Ub-specific processing protease (also called ubiquitin hydrolase) (Miller *et al.*, 1989) and its mammalian homologs (Wilkinson *et al.*, 1989).

**Substrate Specificity of the *UBP1* Protease**—Extracts from *E. coli* carrying either the pUC19 vector alone or the pUC19-based pJT70 plasmid that expressed the yeast *UBP1* protease were added to a variety of Ub-containing test proteins (see "Experimental Procedures"). The reaction products were fractionated by SDS-PAGE, followed either by fluorography (for

the purified <sup>35</sup>S-labeled test proteins) (Fig. 4, A-C) or by immunoblotting with anti-Ub antibodies (Fig. 4D).

The *UBP1* protease was ATP-independent (data not shown), and cleaved ubiquitin from the tested carboxyl-terminal extensions of ubiquitin, such as Ub-X-βgal or Ub-X-DHFR, where X is any of the 20 amino acid residues except proline (Figs. 1, 4A, and data not shown). When the residue X was proline, the *UBP1*-mediated deubiquitination was strongly inhibited (Fig. 4A). This pattern of substrate specificity was indistinguishable from that observed with *S. cerevisiae* using the same test proteins either *in vivo* (Bachmair *et al.*, 1986; Bachmair and Varshavsky, 1989) or in whole cell extracts (Fig. 1 and data not shown).

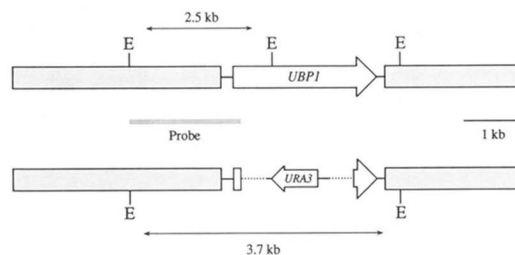
*UBP1* could also deubiquitinate natural ubiquitin fusions such as UBI1-UBI3, which are precursors of both ubiquitin and specific ribosomal proteins (see Introduction). Deubiquitination of UBI2 and UBI3 by the *UBP1* protease (the amino acid sequence of UBI1 is identical to that of UBI2 (Özkaynak *et al.*, 1987)) was demonstrated by mixing extracts of *E. coli* that either contained or lacked *UBP1* with *E. coli* extracts containing UBI2 or UBI3. Deubiquitination of the test substrates was monitored by SDS-PAGE and immunoblotting



**FIG. 4. Substrate specificity of the UBP1 protease.** A, purified,  $^{35}$ S-labeled Ub-M- $\beta$ gal (Fig. 1A, V) and Ub-P- $\beta$ gal (the latter containing proline instead of methionine at the Ub- $\beta$ gal junction) were incubated with UBP1-containing or UBP1-lacking *E. coli* extracts, or with a *S. cerevisiae* extract, as indicated on top of the lanes, followed by electrophoretic analysis of  $\beta$ gal. The bands of Ub-X- $\beta$ gal and X- $\beta$ gal are indicated on the left. B, kinetics of *in vitro* deubiquitination of the  $^{35}$ S-labeled Ub-M- $\beta$ gal (Fig. 1A, V) by extracts from *S. cerevisiae* carrying either the wild-type or disruption/deletion (*ubp1-Δ1*) alleles of the yeast *UBP1* gene. After incubations for specified periods of time,  $\beta$ -galactosidase was analyzed by SDS-PAGE. C, *in vivo* deubiquitination of DHFR-Ub-M- $\beta$ gal (Fig. 1A, VI) in [ $^{35}$ S]methionine-labeled *E. coli* expressing the yeast *UBP1* gene. Lane a, *E. coli* expressing both UBP1 and Ub-M- $\beta$ gal (Fig. 1A, V) were labeled with [ $^{35}$ S]methionine, followed by extraction, immunoprecipitation, and electrophoretic analysis of  $\beta$ -galactosidase. Lane b, same as lane a but with *E. coli* lacking the yeast *UBP1* gene and expressing DHFR-Ub-M- $\beta$ gal. Lane c, same as lane b but with *E. coli* expressing both UBP1 and DHFR-Ub-M- $\beta$ gal. Arrowheads indicate the species which are either partial (apparently *in vivo*) degradation products of DHFR-Ub-M- $\beta$ gal, with cleavages in its DHFR moiety, or the primary translation products due to initiations within the DHFR-coding region of the ORF for DHFR-Ub-M- $\beta$ gal (all of these species were converted into M- $\beta$ gal in the presence of the UBP1 protease; lane c). D, *in vitro* deubiquitination assays with natural ubiquitin fusions. Lane a, an extract from *E. coli* expressing *UBI2* (Fig. 1A, I) was mixed with an extract from *E. coli* lacking the yeast UBP1 protease, followed by an incubation, SDS-PAGE, and detection of ubiquitin-containing species using immunoblot analysis with an anti-Ub antibody (see "Experimental Procedures"). Lane b, same as lane a but with the second extract being from yeast. Lane c, same as lane b but with the second extract being from *E. coli* expressing *UBP1*. Lane d, same as lane a but with the first extract being from *E. coli* expressing *UBI3* (Fig. 1A, II). Lane e, same as lane d but with the second extract being from yeast. Lane f, same as lane e but with the second extract being from *E. coli* expressing *UBP1*. Note that UBP1 cleaves ubiquitin from both *UBI3* and its partial degradation products seen in c. Lane g, same as lane a, but with the first extract being from *E. coli* expressing *UBI4* (Fig. 1A, III). Lane h, same as lane g but with the second extract being from yeast. Lane i, same as lane h but with the second extract being from *E. coli* expressing *UBP1*. Partially deubiquitinated derivatives of *UBI4* (which is denoted as Ub<sub>5</sub>) are indicated on the right. (See "Experimental Procedures" for details.)

with anti-Ub antibodies (Fig. 4D).

Polyubiquitin, a naturally occurring spacerless, head-to-tail fusion of several ubiquitin moieties (Özkaynak *et al.*, 1984, 1987; Dworkin-Rastl *et al.*, 1984; Schlesinger and Bond, 1987), is the only linear ubiquitin fusion that was not cleaved in a UBP1-containing *E. coli* extract; however, the same yeast polyubiquitin (encoded by *UBI4* (Özkaynak *et al.*, 1987)) was converted into mature ubiquitin by a whole-cell yeast extract (Fig. 4D), indicating that *S. cerevisiae* contains Ub-specific



**FIG. 5. Wild-type and null alleles of the UBP1 gene.** Replacement of the wild-type *UBP1* allele with the deletion/disruption *ubp1-Δ1::URA3* allele generates a new arrangement of *EcoRI* sites detectable by Southern hybridization analysis of Ura<sup>+</sup> *S. cerevisiae* transformants using the indicated DNA probe. (See "Experimental Procedures" for details.)

proteases whose *in vitro* substrate specificity is different from those of UBP1 and YUH1. More recent data<sup>3</sup> indicate that UBP1, while inactive with polyubiquitin *in vitro* (in *E. coli* extracts), was able to cleave it *in vivo*, when UBP1 and polyubiquitin were expressed together in *E. coli* (see "Discussion").

The UBP1 protease also did not cleave *in vitro* a branched ubiquitin conjugate such as the multiubiquitin chain (see Introduction). Specifically, no cleavage of the *in vitro*-formed,  $\beta$ -galactosidase-attached multiubiquitin chain (Chau *et al.*, 1989) by the UBP1 protease was observed when the UBP1-containing *E. coli* extract was added to  $^{35}$ S-labeled, multiubiquitinated  $\beta$ -galactosidase (data not shown).

The UBP1 protease deubiquitinated carboxyl-terminal ubiquitin fusions even if the ubiquitin moiety bore in addition an amino-terminal extension. This was shown by expressing a fusion protein DHFR-Ub-M- $\beta$ gal (Fig. 1A, VI) in *E. coli* either by itself or together with UBP1 (see "Experimental Procedures"). The cells were labeled with [ $^{35}$ S]methionine, extracts were prepared, and immunoprecipitated with a monoclonal antibody to  $\beta$ -galactosidase (Fig. 4C). When *E. coli* did not carry a *UBP1*-expressing plasmid, several protein species, including the one whose apparent molecular mass equaled that of DHFR-Ub-M- $\beta$ gal, were precipitated by the anti- $\beta$ gal antibody (Fig. 4C, lane b). However, in *E. coli* that expressed both DHFR-Ub-M- $\beta$ gal and UBP1, all of these  $\beta$ -galactosidase species were converted to the one that comigrated with M- $\beta$ gal (Fig. 4C, lane c). Note that smaller  $\beta$ -galactosidase-containing proteins (Fig. 4C, arrowheads) in *E. coli* lacking UBP1 and expressing DHFR-Ub-M- $\beta$ gal were due either to alternative translation start sites or proteolytic cleavages within the DHFR moiety of the fusion protein, inasmuch as all of these species were converted by UBP1 into M- $\beta$ gal (Fig. 4C, lane c).

**Null *ubp1* Mutants**—Haploid *S. cerevisiae* strains in which 81% of the *UBP1* ORF was deleted and replaced with the *URA3* gene were constructed as described under "Experimental Procedures" (Fig. 5). These *ubp1-Δ1* strains grew normally and lacked obvious phenotypic abnormalities (data not shown). In addition, extracts from *ubp1-Δ1* cells were indistinguishable from wild-type extracts in their ability to deubiquitinate ubiquitin fusions such as Ub-M-DHFR and Ub-M- $\beta$ gal (Fig. 4B and data not shown), indicating that the family of Ub-specific proteases in yeast is not limited to UBP1 and YUH1.

## DISCUSSION

We have cloned a gene for UBP1, a new member of the family of Ub-specific processing proteases in *S. cerevisiae*.

<sup>3</sup> R. Baker and A. Varshavsky, manuscript in preparation.

The UBP1 protease is an 809-residue protein that has no significant sequence similarities to known proteins, including the only other previously cloned yeast Ub-specific processing protease, the 236-residue YUH1 (Miller *et al.*, 1989). While the YUH1 protease is virtually inactive against relatively long ubiquitin fusions such as Ub- $\beta$ gal (Miller *et al.*, 1989),<sup>2</sup> the UBP1 protease of the present work cleaves at the carboxyl terminus of ubiquitin in natural and engineered fusion proteins irrespective of their size or the presence of an amino-terminal ubiquitin extension (see "Results"). The latter property of UBP1 makes it possible to express ubiquitin fusions to proteins of interest as "triple" fusions, whose amino-terminal ubiquitin extension, *e.g.* the biotin-binding protein streptavidin, could be used to purify the fusion by affinity chromatography (from bacteria such as *E. coli* that lack Ub-specific proteases) before the UBP1-mediated deubiquitination step.

Among the ubiquitin fusions tested in the present work, polyubiquitin is the only linear fusion that is not a substrate for the UBP1 protease *in vitro* (see "Results"). Since the amino-terminal region of ubiquitin, and in particular its amino-terminal methionine residue, is an integral part of the compact ubiquitin globule (Vijay-Kumar *et al.*, 1987; Wilkinson, 1988), the cleavable Gly-Met bond between two adjacent ubiquitin moieties in polyubiquitin may be sterically shielded to a greater extent than are analogous Gly-X bonds in ubiquitin fusions that lack the Ub-Ub motif. It has recently been found that UBP1, while unable to deubiquitinate polyubiquitin *in vitro*, can do so *in vivo*, when coexpressed with polyubiquitin in *E. coli*.<sup>3</sup> A likely explanation is that the UBP1-mediated deubiquitination of polyubiquitin *in vivo* is cotranslational or, more generally, occurs before the folding of the newly formed, multidomain polyubiquitin protein is complete.

UBP1 also did not cleave *in vitro* a branched ubiquitin conjugate such as the multiubiquitin chain (see Introduction and "Results"). It remains to be determined whether UBP1 fails to cleave the multiubiquitin chain *in vivo* as well.

Although UBP1 is able to cleave all of the tested natural and engineered ubiquitin fusions at least *in vivo* (see above), and although the processing of natural ubiquitin precursors is essential for cell viability (Finley *et al.*, 1989), *UBP1* is a nonessential gene: the null *ubp1* mutant is viable, grows at a wild-type rate, and is otherwise phenotypically normal (see "Results"). Moreover, the double *ubp1 yuh1* mutant, which lacks both the UBP1 and YUH1 proteases, is also phenotypically normal, and retains the ability to deubiquitinate both linear and branched Ub-protein conjugates.<sup>3</sup> The multiplicity of functionally overlapping Ub-specific processing proteases in *S. cerevisiae* accounts, in hindsight, for the failure of our earlier efforts to clone the genes for these proteases using yeast-based genetic screens. While the strategy of sib selection and functional cloning in *E. coli* that yielded the *UBP1* gene bypassed the problem of gene multiplicity, it was initially viewed as a strategy unlikely to succeed for other reasons (see "Results"). More recent work,<sup>3</sup> using another *E. coli*-based genetic screen, has yielded two more *S. cerevisiae* genes for Ub-specific processing proteases, named *UBP2* and *UBP3*. Thus, the yeast *UBP* gene family contains at least four genes, including *YUH1*; a variety of evidence indicates that other eukaryotes have at least as many (Rose, 1988; Wilkinson *et al.*, 1989; Jonnalagadda *et al.*, 1989). The continuing genetic and biochemical analyses of the single and multiple null *ubp* mutants should allow more precise definitions of the apparently overlapping functions of these proteases.

Prior to the cloning of *YUH1* and *UBP1*, the presence of UB-specific processing proteases in *S. cerevisiae* and other

eukaryotes was inferred from the existence of polyubiquitin genes (Özkaynak *et al.*, 1984, 1987; Dworkin-Rastl *et al.*, 1984), and from the demonstration that engineered ubiquitin fusions such as Ub-X- $\beta$ gals were precisely deubiquitinated *in vivo* or in cell-free extracts (Bachmair *et al.*, 1986; Gonda *et al.*, 1989). Another finding, that Ub-X- $\beta$ gals were deubiquitinated irrespective of the identity of a residue X at the Ub- $\beta$ gal junction, led to the discovery of the N-end rule, which relates the *in vivo* half-life of a protein to the identity of its amino-terminal residue (Bachmair *et al.*, 1986). Distinct versions of the N-end rule have since been shown to operate in all eukaryotes examined, from yeast to mammals (Bachmair *et al.*, 1986; Varshavsky *et al.*, 1988; Reiss *et al.*, 1988; Bachmair and Varshavsky, 1989; Gonda *et al.*, 1989; Johnson *et al.*, 1990; Bartel *et al.*, 1990). However, since ubiquitin and Ub-specific enzymes are absent from bacteria (Finley *et al.*, 1988), the ubiquitin fusion technique could not be used to determine whether the N-end rule operates in prokaryotes as well. The isolation of *UBP1* has allowed us to bypass this difficulty, as *E. coli* transformed with a *UBP1*-expressing plasmid acquires the ability to deubiquitinate Ub-X- $\beta$ gals irrespective of the identity of the junctional residue X, proline being the single exception. Measurements of the *in vivo* half-lives of the resultant X- $\beta$ gals in *E. coli* have shown that a distinct version of the N-end rule operates in bacteria as well.<sup>4</sup>

Availability of the *UBP1* gene and the purified UBP1 protease will increase the versatility of the ubiquitin fusion-based technique for the *in vivo* or *in vitro* generation of proteins (including short polypeptides) that bear predetermined residues at their amino termini, a method with applications in both basic research (Baker and Varshavsky, 1991; Bartel *et al.*, 1990; Johnson *et al.*, 1990; Townsend *et al.*, 1988) and biotechnology (Lu *et al.*, 1990; Butt *et al.*, 1989; Sabin *et al.*, 1989; Ecker *et al.*, 1989; Yoo *et al.*, 1989; Mak *et al.*, 1989).

**Acknowledgments**—We thank Vincent Chau for the anti-Ub antibody and Rohan Baker for the pRBU1 plasmid. We also thank Erica Johnson, Tom Shrader, and especially Rohan Baker for comments on the manuscript and Barbara Doran for secretarial assistance.

## REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds) (1989) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.
- Bachmair, A., and Varshavsky, A. (1989) *Cell* **56**, 1019–1032.
- Bachmair, A., Finley, D., and Varshavsky, A. (1986) *Science* **234**, 179–186.
- Baker, R. T., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 1090–1094.
- Ball, E., Karlik, C. C., Beall, C. J., Saville, D. L., Sparrow, J. C., Bullard, B., and Fyrberg, E. A. (1987) *Cell* **51**, 221–228.
- Bartel, B., Wüning, I., and Varshavsky, A. (1990) *EMBO J.* **9**, 3179–3189.
- Bonner, W., Hatch, C. L., and Wu, R. S. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 157–172, Plenum Press, New York.
- Botstein, D., Falco, S. C., Stewart, S. E., Brennan, M., Scherer, S., Stinchcomb, D. T., Struhl, K., and Davis, R. W. (1979) *Gene (Amst.)* **8**, 17–24.
- Boyer, H. W., and Roulland-Dussoix, D. (1969) *J. Mol. Biol.* **41**, 459–472.
- Butt, T., Jonnalagadda, S., Monia, B. P., Sternberg, E. J., Marsh, J. A., Stadel, J. M., Ecker, D. J., and Crooke, S. T. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 2540–2544.
- Chau, V., Tobias, J. W., Bachmair, A., Marriotti, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989) *Science* **243**, 1576–1583.
- Ciechanover, A., and Schwartz, A. L. (1989) *Trends Biochem. Sci.* **14**, 483–488.
- Dale, R. M. K., McClure, B. A., and Houchins, J. P. (1985) *Plasmid* **13**, 31–40.

<sup>4</sup> J. Tobias, T. Shrader, G. Rocap, and A. Varshavsky, manuscript in preparation.

- Dworkin-Rastl, E., Shrutkowski, A., and Dworkin, M. B. (1984) *Cell* **39**, 321-325
- Ecker, D. J., Stadel, J. M., Butt, R. T., Marsha, J. A., Monia, B. P., Powers, D. A., Gorman, J. A., Clark, P. E., Warren, F., Shatzman, A., and Crooke, S. T. (1989) *J. Biol. Chem.* **264**, 7715-7719
- Finley, D., Ciechanover, A., and Varshavsky, A. (1984) *Cell* **37**, 43-55
- Finley, D., Özkaynak, E., and Varshavsky, A. (1987) *Cell* **48**, 1035-1046
- Finley, D., Özkaynak, E., Jentsch, S., McGrath, J. P., Bartel, B., Pazin, M., Snapka, R. M., and Varshavsky, A. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 39-75, Plenum Press, New York
- Finley, D., Bartel, B., and Varshavsky, A. (1989) *Nature* **338**, 394-401
- Glötzer, M., Murray, A., and Kirschner, M. (1991) *Nature* **349**, 132-138
- Gonda, D. K., Bachmair, A., Wüning, L., Tobias, J. W., Lane, W. S., and Varshavsky, A. (1989) *J. Biol. Chem.* **264**, 16700-16712
- Gregori, L., Poosch, M. S., Cousins, G., and Chau, V. (1990) *J. Biol. Chem.* **265**, 8354-8357
- Guarente, L., Yocum, R. R., and Gillard, R. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 7410-7414
- Hershko, A. (1988) *J. Biol. Chem.* **263**, 15237-15240
- Hochstrasser, M., Ellison, M., Chau, V., and Varshavsky, A. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 4606-4610
- Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) *J. Bacteriol.* **153**, 163-168
- Jabben, M., Shanklin, J., and Vierstra, R. D. (1989) *J. Biol. Chem.* **264**, 4998-5005
- Jentsch, S., McGrath, J. P., and Varshavsky, A. (1987) *Nature* **329**, 131-134
- Jentsch, S., Seufert, W., Sommer, T., and Reins, H.-A. (1990) *Trends Biochem. Sci.* **15**, 195-198
- Johnson, E. S., Gonda, D. K., and Varshavsky, A. (1990) *Nature* **346**, 287-291
- Jonnalagadda, S., Butt, T. R., Monia, B. P., Mirabelli, C. K., Gotlib, L., Ecker, D. J., and Crooke, S. T. (1989) *J. Biol. Chem.* **264**, 10637-10642
- Kaufman, B. T. (1974) *Methods Enzymol.* **34**, 272-281
- Lederberg, J. (1989) *Genetics* **121**, 395-399
- Leung, D. W., Spencer, S. A., Cachianes, G., Hammonds, R. G., Collins, C., Henzel, W. J., Barnard, R., Waters, M. J., and Wood, W. I. (1987) *Nature* **330**, 537-543
- Lowe, J., Blanchard, A., Morrel, K., Lennox, G., Reynolds, L., Billett, M., Landon, M., and Mayer, R. J. (1988) *J. Pathol.* **155**, 9-15
- Lu, C., Yang, Y.-F., Ohashi, H., and Walfish, P. G. (1990) *Biochem. Biophys. Res. Commun.* **171**, 138-142
- Mak, P., McDonnell, D. P., Weigel, N. L., Schrader, W. T., and O'Malley, B. W. (1989) *J. Biol. Chem.* **264**, 21613-21618
- Matsui, S., Sandberg, A. A., Negoro, S., Seon, B. K., and Goldstein, G. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 1535-1539
- Mayer, A. N., and Wilkinson, K. D. (1989) *Biochemistry* **28**, 166-172
- McCormick, M. (1987) *Methods Enzymol.* **151**, 445-449
- Miller, H. I., Henzel, W. J., Ridgway, J. B., Kuang, W.-J., Chisholm, V., and Liu, C.-C. (1989) *Bio/Technology* **7**, 698-704
- Monia, B. P., Ecker, D. J., and Crooke, S. P. (1990) *Bio/Technology* **8**, 209-215
- Olson, T. S., and Dice, J. F. (1989) *Curr. Opin. Cell Biol.* **1**, 1194-1200
- Özkaynak, E., Finley, D., and Varshavsky, A. (1984) *Nature* **312**, 663-666
- Özkaynak, E., Finley, D., Solomon, M. J., and Varshavsky, A. (1987) *EMBO J.* **6**, 1429-1439
- Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2444-2448
- Perry, G., Friedman, R., Shaw, G., and Chau, V. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 3033-3036
- Pickart, C. M., and Rose, I. A. (1985) *J. Biol. Chem.* **260**, 1573-1581
- Pickart, C. M. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 77-100, Plenum Press, New York
- Reiss, Y., Kaim, D., and Hershko, A. (1988) *J. Biol. Chem.* **263**, 2693-2698
- Reynolds, P., Koken, M. H. M., Hoeijmakers, J. H. J., Prakash, S., and Prakash, L. (1990) *EMBO J.* **9**, 1423-1430
- Rose, I. A. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 135-155, Plenum Press, New York
- Rose, M. D., Novick, P., Thomas, J. H., Botstein, D., and Fink, G. R. (1987) *Gene (Amst.)* **60**, 237-243
- Rothstein, R. (1983) *Methods Enzymol.* **101**, 202-211
- Sabin, E. A., Lee-Ng, C. T., Shuster, J. R., and Barr, P. J. (1989) *Bio/Technology* **7**, 705-709
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Schlesinger, M., and Bond, U. (1987) *Oxf. Surv. Eukaryotic Genes* **4**, 77-91
- Schneider, R., Eckerskom, C., Lottspeich, F., and Schweiger, M. (1990) *EMBO J.* **9**, 1431-1437
- Sharp, P. M., and Li, W.-H. (1987) *Nucleic Acids Res.* **15**, 1281-1295
- Sherman, F., Fink, G. R., and Hicks, J. B. (1986) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Siegelman, M., Bond, M. W., Gallatin, W. M., St. John, T., Smith, H. T., Fried, V. A., and Weissman, I. L. (1986) *Science* **231**, 823-829
- Sullivan, M. L., and Vierstra, R. D. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 9861-9865
- Townsend, A., Bastin, J., Gould, K., Brownlee, G., Andrew, M., Coupar, B., Boyle, D., Chan, S., and Smith, G. (1988) *J. Exp. Med.* **168**, 1211-1224
- Varshavsky, A., Bachmair, A., Finley, D., Gonda, D., and Wüning, I. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 287-321, Plenum Press, New York
- Vijay-Kumar, S., Bugg, C. E., and Cook, W. J. (1987) *J. Mol. Biol.* **194**, 531-544
- Wilkinson, K. D. (1988) in *Ubiquitin* (Rechsteiner, M., ed) pp. 5-38, Plenum Press, New York
- Wilkinson, K. D., Lee, K., Deshpande, S., Duerksen-Hughes, P., Boss, J. M., and Pohl, J. (1989) *Science* **246**, 670-673
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene (Amst.)* **33**, 103-119
- Yoo, Y., Rote, K., and Rechsteiner, M. (1989) *J. Biol. Chem.* **264**, 17078-17083